

Optimal Average Satisfaction and Extended Justified Representation in Polynomial Time

Piotr Skowron¹, Martin Lackner², Edith Elkind², and Luis Sánchez-Fernández³

¹TU Berlin, Germany ²University of Oxford, UK

³Universidad Carlos III de Madrid, Spain

Abstract. In this short note, we describe an approval-based committee selection rule that admits a polynomial-time algorithm and satisfies the Extended Justified Representation (EJR) axiom. This rule is based on approximately maximizing the PAV score, by means of local search. Our proof strategy is to show that this rule provides almost optimal average satisfaction to all cohesive groups of voters, and that high average satisfaction for cohesive groups implies extended justified representation.

1 Introduction

Committee selection rules, i.e., rules that, given a collection of voters' preferences over candidates, output a fixed-size set of winners (a committee), have received a considerable amount of attention in the last few years [3]. For approval-based committee selection rules, i.e., rules where each voter reports a subset of candidates that she approves of, an influential recent paper of Aziz et al. [1] has proposed an axiom called *extended justified representation (EJR)*. Informally, this axiom requires that if there is a large group of voters whose preferences have substantial overlap, then this group should be well-represented in the committee. Aziz et al. [1] have established that, among existing committee selection rules, Proportional Approval Voting (PAV) is the only rule that satisfies this axiom. Unfortunately, computing the output of this rule is NP-hard [2], and, moreover, Aziz et al. [1] have established that verifying if a given committee provides EJR is computationally hard as well. Therefore, it was conjectured that finding committees that provide EJR is NP-hard.

In this paper, we show that this conjecture is false: we describe a polynomial-time procedure that is guaranteed to output a committee that provides EJR. Our proof makes use of a recently introduced notion of *average satisfaction* [4,6]: we show that our rule offers very high average satisfaction to all groups of voters that are large and cohesive, and deduce from this that it satisfies EJR. Thus, in a sense, the welfare guarantees provided by our rule are even much stronger than those offered by EJR.

2 Preliminaries

An *election* is a pair $E = (N, C)$, where $N = \{1, \dots, n\}$ is a set of *voters* and $C = \{c_1, \dots, c_m\}$ is a set of *candidates*. Voters in N have *approval preferences*:

each voter in N approves a subset of candidates in C . For each $i \in N$ we denote by $A_i \subseteq C$ the set of candidates approved by voter i . We are interested in procedures that, given an election E and a positive integer k , $1 \leq k \leq |C|$, output a non-empty collection of subsets of candidates $W \subseteq C$ of size exactly k ; such procedures are called *committee selection rules*.

Given an election $E = (N, C)$, we define the PAV-score of a committee $W \subseteq C$ as

$$\text{pav-sc}(W) = \sum_{i=1}^n \sum_{j=1}^{|A_i \cap W|} \frac{1}{j}.$$

Proportional Approval Voting (PAV) is the committee selection rule that, given an election E and a committee size k , outputs a committee of size W with the highest PAV-score; finding a committee in the output of this rule is NP-hard [2,5].

We say that a group of voters $V \subseteq N$ is ℓ -cohesive if $|V| \geq \lceil \ell \cdot \frac{n}{k} \rceil$ and $|\bigcap_{i \in V} A_i| \geq \ell$.

Definition 1 (Sánchez-Fernández et al. [4]). *Given a committee W , the average satisfaction of a group of voters $V \subseteq N$ with respect to W is defined as*

$$\text{avsw}(V) = \frac{1}{|V|} \sum_{i \in V} |A_i \cap W|.$$

Definition 2 (Aziz et al. [1]). *A committee W provides extended justified representation (EJR) if for every $\ell > 0$ and every ℓ -cohesive group of voters V there exists a voter $i \in V$ who approves at least ℓ members of W , i.e., $|A_i \cap W| \geq \ell$. A committee selection rule satisfies extended justified representation if for every election $E = (N, C)$ and every k with $1 \leq k \leq |C|$ all committees that it outputs provide EJR.*

Sánchez-Fernández et al. [4] show that if a committee W provides EJR then for every $\ell > 0$ and every ℓ -cohesive group of voters V it holds that the average satisfaction of V with respect to W is at least $\frac{\ell-1}{2}$. We observe that, conversely, if a committee provides very high average satisfaction to all cohesive groups then it provides EJR.

Lemma 3. *Consider an election $E = (N, C)$ and a committee W . If for every ℓ -cohesive group it holds that its average satisfaction with respect to W is strictly greater than $\ell - 1$, then W provides EJR.*

Proof. Let V be an ℓ -cohesive group. Since the average satisfaction of voters in V is greater than $\ell - 1$, there is at least one voter $i \in V$ with $|A_i \cap W| \geq \ell$. \square

3 Local search Algorithm for PAV

We define a new rule, LS-PAV, which is a local search algorithm for PAV (see Figure 1).

Algorithm 1: LS-PAV: a local search algorithm for PAV

$W \leftarrow k$ arbitrary candidates from C
while there exist $c \in W$ and $c' \in C \setminus W$ such that
 $\text{pav-sc}((W \setminus \{c\}) \cup \{c'\}) \geq \text{pav-sc}(W) + \frac{n}{k^2}$ **do**
 $W \leftarrow (W \setminus \{c\}) \cup \{c'\}$
return W

Theorem 4. Consider an election $E = (N, C)$ and a positive integer k , $1 \leq k \leq |C|$. Let W be a winning committee chosen by LS-PAV on (E, k) . Then for every $\ell > 0$ and every ℓ -cohesive group V it holds that $\text{avs}_W(V) > \ell - 1$, i.e., all ℓ -cohesive groups have average satisfaction of more than $\ell - 1$.

Proof. Assume for the sake of contradiction that for some (E, k) LS-PAV outputs a committee W such that there exists an ℓ -cohesive group V with $\text{avs}_W(V) \leq \ell - 1$. Let $w_i = |W \cap A_i|$.

As V is ℓ -cohesive, there exist ℓ candidates approved by all voters in V . At least one such candidate does not appear in W , since otherwise we would have $\text{avs}_W(V) \geq \ell$. Let c be some such candidate. Now consider a candidate $c' \in W$. If we remove c' from the committee and add c instead, we increase the PAV-score of W by

$$\begin{aligned}
\Delta(c, c') &\geq \underbrace{\sum_{i \in V: c' \notin A_i} \frac{1}{w_i + 1}}_{\text{adding } c} - \underbrace{\sum_{i \in N \setminus V: c' \in A_i} \frac{1}{w_i}}_{\text{removing } c'} \\
&= \sum_{i \in V} \frac{1}{w_i + 1} - \sum_{i \in N: c' \in A_i} \frac{1}{w_i} + \sum_{i \in V: c' \in A_i} \left(\frac{1}{w_i} - \frac{1}{w_i + 1} \right).
\end{aligned}$$

Note that $\Delta(c, c')$ may be negative for some $c' \in W$. By the inequality between arithmetic and harmonic means we obtain

$$\sum_{i \in V} \frac{1}{w_i + 1} \geq \frac{|V|^2}{\sum_{i \in V} (w_i + 1)} = \frac{|V|}{\frac{\sum_{i \in V} w_i}{|V|} + 1} = \frac{|V|}{\text{avs}_W(V) + 1} \geq \frac{|V|}{\ell}. \quad (1)$$

Now, observe that

$$\begin{aligned}
\sum_{c' \in W} \Delta(c, c') &\geq \sum_{c' \in W} \left(\sum_{i \in V} \frac{1}{w_i + 1} - \sum_{i \in N: c' \in A_i} \frac{1}{w_i} + \sum_{i \in V: c' \in A_i} \left(\frac{1}{w_i} - \frac{1}{w_i + 1} \right) \right) \\
&= k \sum_{i \in V} \frac{1}{w_i + 1} - \sum_{c' \in W} \sum_{i \in N: c' \in A_i} \frac{1}{w_i} + \sum_{c' \in W} \sum_{i \in V: c' \in A_i} \left(\frac{1}{w_i} - \frac{1}{w_i + 1} \right) \\
&= k \sum_{i \in V} \frac{1}{w_i + 1} - \sum_{i \in N} \sum_{c' \in W \cap A_i} \frac{1}{w_i} + \sum_{i \in V} \sum_{c' \in W \cap A_i} \left(\frac{1}{w_i} - \frac{1}{w_i + 1} \right) \\
&= k \sum_{i \in V} \frac{1}{w_i + 1} - n + |V| - \sum_{i \in V} \frac{w_i}{w_i + 1}
\end{aligned}$$

$$\begin{aligned}
&= k \sum_{i \in V} \frac{1}{w_i + 1} - n + |V| - \sum_{i \in V} \left(1 - \frac{1}{w_i + 1}\right) \\
&= (k + 1) \sum_{i \in V} \frac{1}{w_i + 1} - n.
\end{aligned}$$

Further, by equation (1) we obtain

$$k \sum_{i \in V} \frac{1}{w_i + 1} - n \geq \frac{k|V|}{\ell} - n \geq \left\lceil \frac{\ell n}{k} \right\rceil \cdot \frac{k}{\ell} - n \geq 0,$$

and hence

$$\sum_{c' \in W} \Delta(c, c') \geq (k + 1) \sum_{i \in V} \frac{1}{w_i + 1} - n \geq \sum_{i \in V} \frac{1}{w_i + 1} \geq \frac{|V|}{\ell} \geq \frac{n}{k}.$$

From the pigeonhole principle it follows that there exists $c' \in W$ such that $\Delta(c, c') \geq \frac{n}{k^2}$, which means that W could not have been returned by our local search algorithm. This completes the proof. \square

Corollary 5. *For PAV all ℓ -cohesive groups have average satisfaction of more than $\ell - 1$.*

Note that Theorem 4 applies not only to LS-PAV but also to PAV, as PAV selects a committee with the maximum PAV-score.

Corollary 6. *LS-PAV satisfies extended justified representation.*

Proof. Let $E = (N, C)$ be an election, let k be a positive integer with $1 \leq k \leq |C|$ and let W be a winning committee chosen by LS-PAV. Further, let V be an ℓ -cohesive group. Then, by Theorem 4 it holds that $\text{avs}_W(V) > \ell - 1$. Consequently, by Lemma 3, W provides EJ. \square

Proposition 7. *LS-PAV is computable in polynomial time.*

Proof. A single improving swap can be clearly found and executed in polynomial time. Now, let us assess how many improvements the local search algorithm may perform. Each improvement increases the total PAV-score of a committee by at least $\frac{n}{k^2}$. The maximum score a committee can get is $n \cdot (1 + 1/2 + \dots + 1/k) = O(n \ln(k))$. Thus, there may be at most $O(k^2 \ln(k))$ improving swaps. \square

Observe that Proposition 7 relies on having a threshold of $\frac{n}{k^2}$ in the definition of the local search algorithm. If we perform a swap each time when $\text{pav-sc}((W \setminus \{c\}) \cup \{c'\}) \geq \text{pav-sc}(W)$, this could potentially lead to a super-polynomial running time, since by a naive argument we could only conclude that each swap increases the total PAV score of a committee by one over the least common multiple of values $1, 2, \dots, k$, which can be exponential with respect to k .

We will now show that the satisfaction guarantee given by Theorem 4 cannot be improved.

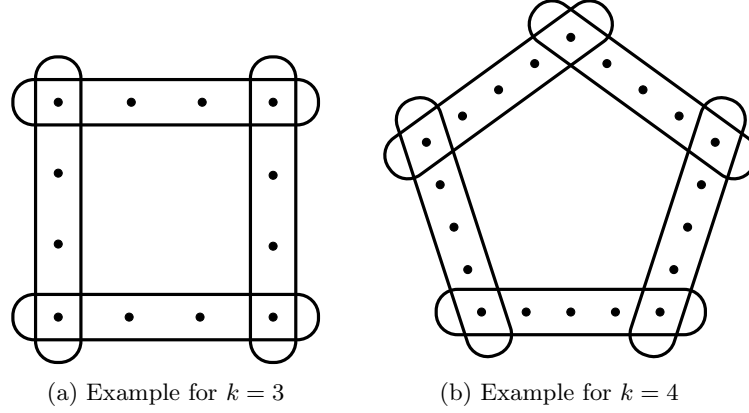


Fig. 1: A visualization of the profiles used in Example 8. Dots represent voters and boxes represent candidates; candidates are approved by those voters that the respective boxes contain.

Example 8. Consider the following election.

$1 \times \{d, a\}$	$2 \times \{a\}$
$1 \times \{a, b\}$	$2 \times \{b\}$
$1 \times \{b, c\}$	$2 \times \{c\}$
$1 \times \{c, d\}$	$2 \times \{d\}$

This profile is schematically shown in Figure 1a. For $k = 3$, we have $\frac{n}{k} = 4$ and consequently all voters that approve a fixed candidate form a 1-cohesive group. The profile is symmetric with respect to candidates so without loss of generality assume that committee $\{a, b, c\}$ is chosen. Let us consider the voters who approve d , i.e., $\{d, a\}$, $2 \times \{d\}$, and $\{c, d\}$. The average satisfaction of this group is $1/2$. Hence we have found a profile where it is impossible to guarantee an average satisfaction for 1-cohesive groups that is better than $1/2$.

Let us now extend this example for $k = 4$, moving from a rectangle shape to a pentagon (see Figure 1b).

$1 \times \{e, a\}$	$3 \times \{a\}$
$1 \times \{a, b\}$	$3 \times \{b\}$
$1 \times \{b, c\}$	$3 \times \{c\}$
$1 \times \{c, d\}$	$3 \times \{d\}$
$1 \times \{d, e\}$	$3 \times \{e\}$

By the same argument as before, we can assume without loss of generality that e is not contained in the committee. The average satisfaction of the respective group is $\frac{2}{5}$ and we see that we cannot guarantee an average satisfaction

of more than $\frac{2}{5}$ for 1-cohesive groups. If we generalize these types of examples to larger k , we can deduce that we cannot guarantee an average satisfaction of more than $\frac{2}{k}$ for 1-cohesive groups. Hence, for arbitrary k , there is no positive constant γ such that we can guarantee an average satisfaction of γ for 1-cohesive groups.

Proposition 9. *Let ℓ be a positive integer and $\gamma > 0$. It is not possible to guarantee an average satisfaction of $\ell - 1 + \gamma$ for ℓ -cohesive groups.*

Proof. We have seen in Example 8 how to construct profiles where 1-cohesive groups cannot have an average satisfaction better than $\frac{2}{k}$. If we choose $k > \frac{2}{\gamma}$, then we have shown the statement for $\ell = 1$. This construction can easily be generalized for arbitrary ℓ . We replace each candidate with ℓ copies; voters approve all copies of previously approved candidates. In this case there is at least one ℓ -cohesive group with only $\ell - 1$ of their joint candidates approved; let this group be V . We have $\text{avs}_W(V) = \ell - 1 + \frac{2}{\gamma}$. As before, for $k > \frac{2}{\gamma}$ we obtain an example showing that an average satisfaction of $\ell - 1 + \gamma$ for ℓ -cohesive groups cannot be guaranteed. \square

References

1. Aziz, H., Brill, M., Conitzer, V., Elkind, E., Freeman, R., Walsh, T.: Justified representation in approval-based committee voting. *Social Choice and Welfare* 48(2), 461–485 (2017)
2. Aziz, H., Gaspers, S., Gudmundsson, J., Mackenzie, S., Mattei, N., Walsh, T.: Computational aspects of multi-winner approval voting. In: 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS). pp. 107–115 (2015)
3. Faliszewski, P., Skowron, P., Slinko, A., Talmon, N.: Multiwinner voting: A new challenge for social choice theory. In: Endriss, U. (ed.) *Trends in Computational Social Choice*. AI Access (2017)
4. Sánchez-Fernández, L., Elkind, E., Lackner, M., Fernández, N., Fisteus, J.A., Banta Val, P., Skowron, P.: Proportional justified representation. In: *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI 2017)*. pp. 670–676. AAAI Press (2017)
5. Skowron, P., Faliszewski, P., Lang, J.: Finding a collective set of items: From proportional multirepresentation to group recommendation. *Artificial Intelligence* 241, 191–216 (2016)
6. Skowron, P., Lackner, M., Brill, M., Peters, D., Elkind, E.: Proportional rankings. CoRR abs/1612.01434 (2016), <http://arxiv.org/abs/1612.01434>